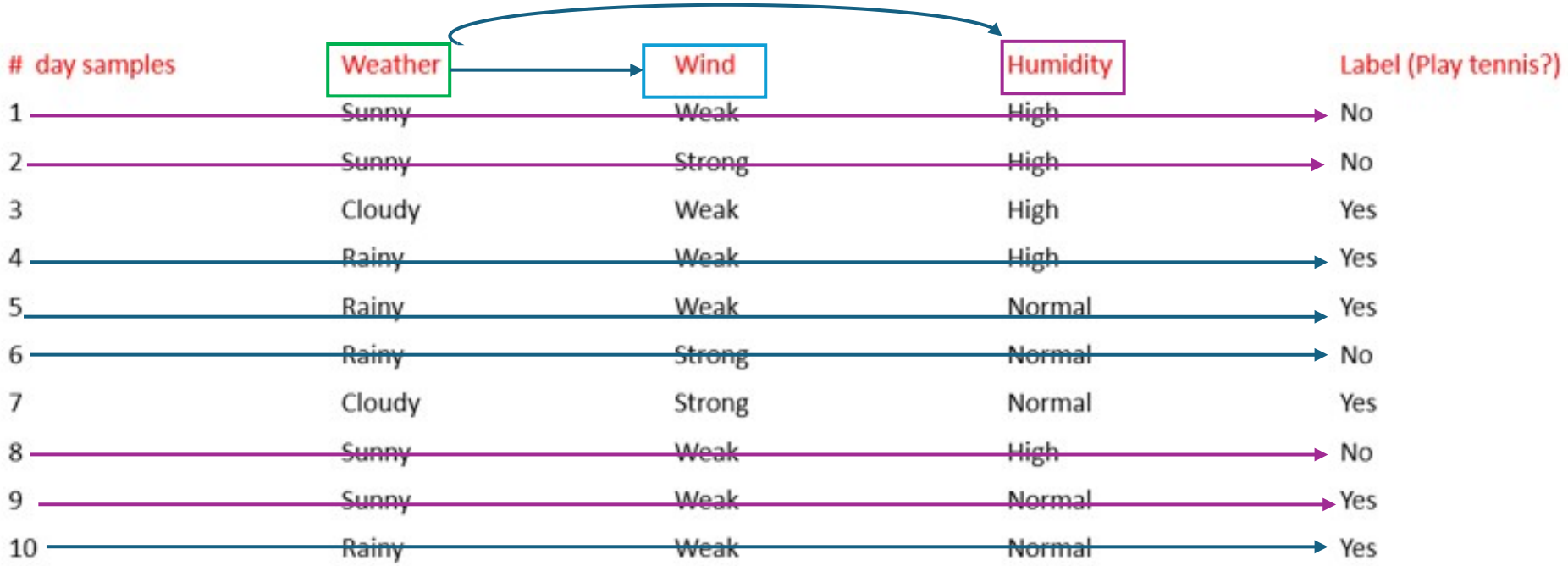
Bayesian Networks

Luis G. Jaimes

Intro (are the features independent ?)

- We already classified this dataset with Naive Bayes. The model worked, but it assumed that Weather, Wind, and Humidity were completely independent given the label. Let's ask whether that assumption is reasonable.
- Think about it intuitively:
- On rainy days, is wind more likely to be strong or weak? 3 out 4
- On sunny days, does humidity tend to be high or low?

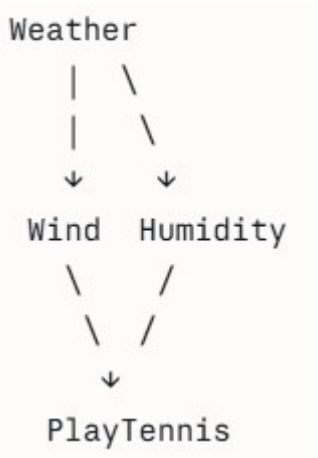


# day samples	Weather	Wind	Humidity	Label (Play tennis?)
1	Sunny	Weak	High	No
2	Sunny	Strong	High	No
3	Cloudy	Weak	High	Yes
4	Rainy	Weak	High	Yes
5	Rainy	Weak	Normal	Yes
6	Rainy	Strong	Normal	No
7	Cloudy	Strong	Normal	Yes
8	Sunny	Weak	High	No
9	Sunny	Weak	Normal	Yes
10	Rainy	Weak	Normal	Yes

Example

Wind and Humidity are also atmospheric; they are influenced by Weather patterns. The decision to play tennis is then a consequence of all three conditions.

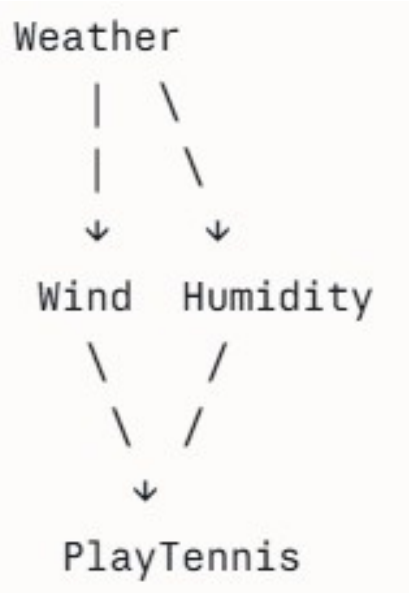
# day samples	Weather	Wind	Humidity	Label (Play tennis?)
1	Sunny	Weak	High	No
2	Sunny	Strong	High	No
3	Cloudy	Weak	High	Yes
4	Rainy	Weak	High	Yes
5	Rainy	Weak	Normal	Yes
6	Rainy	Strong	Normal	No
7	Cloudy	Strong	Normal	Yes
8	Sunny	Weak	High	No
9	Sunny	Weak	Normal	Yes
10	Rainy	Weak	Normal	Yes



Reading this graph:

- Weather directly influences Wind and Humidity
- Weather, Wind, and Humidity all directly influence PlayTennis
- Wind and Humidity do not influence each other directly, any correlation between them is explained by their shared parent Weather

Steps

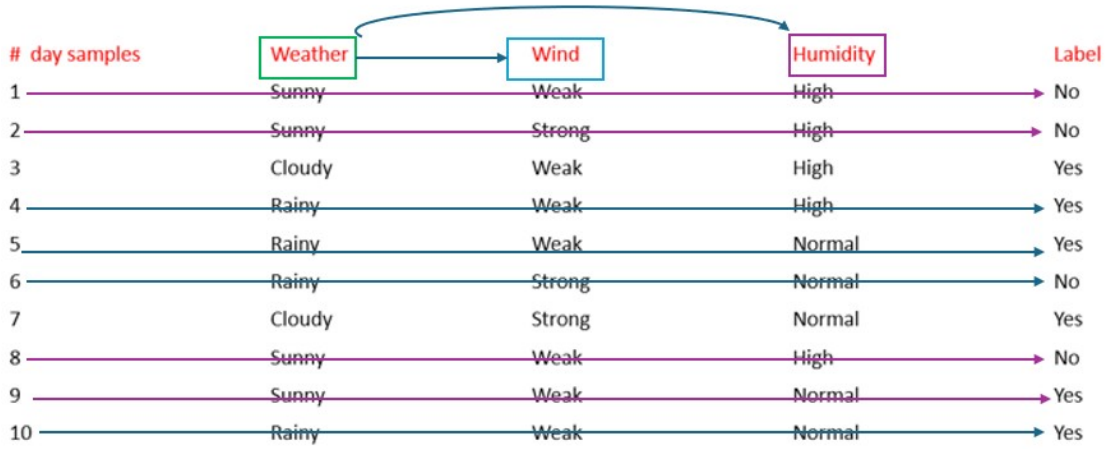


- Step 1: Think about the structure before touching the da
- Step 2: Get the priors for root nodes (CPT for Weather)
- Step 3: Build CPT for Wind given Weather P(wind | weather)
- Step 4: Build CPT for Humidity given Weather
- Step 5: Build CPT for PlayTennis given Weather, Wind, Humidity

Steps

Step 2: Priors for the root, CPTs for weather

Weather	P
Sunny	4/10 = 0.40
Cloudy	2/10 = 0.20
Rainy	4/10 = 0.40



# day samples	Weather	Wind	Humidity	Label
1	Sunny	Weak	High	No
2	Sunny	Strong	High	No
3	Cloudy	Weak	High	Yes
4	Rainy	Weak	High	Yes
5	Rainy	Weak	Normal	Yes
6	Rainy	Strong	Normal	No
7	Cloudy	Strong	Normal	Yes
8	Sunny	Weak	High	No
9	Sunny	Weak	Normal	Yes
10	Rainy	Weak	Normal	Yes

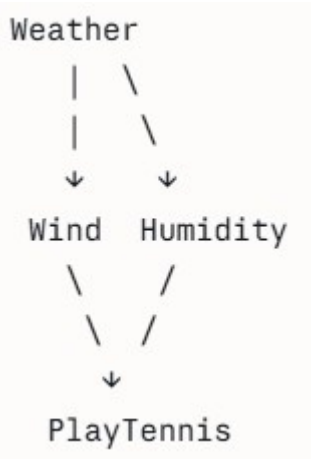
Step 3: CPT for wind given weather P(wind | weather)

Weather	P(Weak)	P(Strong)
Sunny	3/4 = 0.75	1/4 = 0.25
Cloudy	1/2 = 0.50	1/2 = 0.50
Rainy	3/4 = 0.75	1/4 = 0.25

When Weather = Sunny (rows 1, 2, 8, 9)
Then, how many out of the sunny are wind = weak (3)

Step4: CPT for wind given weather P(humidity | weather)

Weather	P(High)	P(Normal)
Sunny	3/4 = 0.75	1/4 = 0.25
Cloudy	1/2 = 0.50	1/2 = 0.50
Rainy	1/4 = 0.25	3/4 = 0.75



Step: P(playtennis | weather, wind, humidity)

Step 5:

Go through every unique combination that appears:

- **Sunny, Weak, High** → rows 1, 8 → both No → P(Yes)=0, P(No)=1
- **Sunny, Strong, High** → row 2 → No → P(Yes)=0, P(No)=1
- **Sunny, Weak, Normal** → row 9 → Yes → P(Yes)=1, P(No)=0
- **Cloudy, Weak, High** → row 3 → Yes → P(Yes)=1, P(No)=0
- **Cloudy, Strong, Normal** → row 7 → Yes → P(Yes)=1, P(No)=0
- **Rainy, Weak, High** → row 4 → Yes → P(Yes)=1, P(No)=0

P(PlayTennis = yes | weather = sunny , wind = weak , humidity = high) □ rows 1, 8 P(NO) =2/2, P(yes)=0

# day samples	Weather	Wind	Humidity	Label (Play tennis?)
1	Sunny	Weak	High	No
2	Sunny	Strong	High	No
3	Cloudy	Weak	High	Yes
4	Rainy	Weak	High	Yes
5	Rainy	Weak	Normal	Yes
6	Rainy	Strong	Normal	No
7	Cloudy	Strong	Normal	Yes
8	Sunny	Weak	High	No
9	Sunny	Weak	Normal	Yes
10	Rainy	Weak	Normal	Yes

Steps

- Step 5: Build CPT for PlayTennis given Weather, Wind, Humidity

  P(PlayTennis | Weather, Wind, Humidity)

Weather	Wind	Humidity	P(Yes)	P(No)
Sunny	Weak	High		
Sunny	Strong	High		
Sunny	Weak	Normal		
Sunny	Strong	Normal		
Cloudy	Weak	High		
Cloudy	Strong	Normal		
Cloudy	Weak	Normal		
Cloudy	Strong	High		
Rainy	Weak	High		
Rainy	Weak	Normal		
Rainy	Strong	Normal		
Rainy	Strong	High		

Steps

- Step 5: Build CPT for PlayTennis given Weather, Wind, Humidity

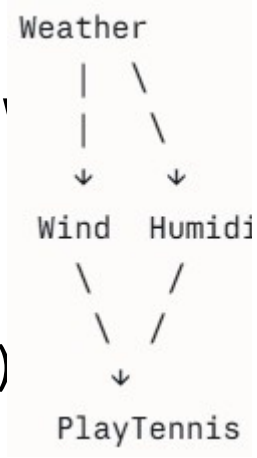
  P(PlayTennis | Weather, Wind, Humidity)

Weather	Wind	Humidity	P(Yes)	P(No)
Sunny	Weak	High	0/2 = 0.00	2/2 = 1.00
Sunny	Strong	High	0/1 = 0.00	1/1 = 1.00
Sunny	Weak	Normal	1/1 = 1.00	0/1 = 0.00
Sunny	Strong	Normal	—	—
Cloudy	Weak	High	1/1 = 1.00	0/1 = 0.00
Cloudy	Strong	Normal	1/1 = 1.00	0/1 = 0.00
Cloudy	Weak	Normal	—	—
Cloudy	Strong	High	—	—
Rainy	Weak	High	1/1 = 1.00	0/1 = 0.00
Rainy	Weak	Normal	2/2 = 1.00	0/2 = 0.00
Rainy	Strong	Normal	0/1 = 0.00	1/1 = 1.00
Rainy	Strong	High	—	—

Getting the priors and conditional probs

Build CPT for Play Tennis given Weather, Wind, Humidity P(play tennis | weather, wind, humidity)

P(W, Wi, H, T) = P(W) × P(Wi|W) × P(H|W) × P(T|



Conditionally Probability Table (CPT) for Weather:

Weather	P
Sunny	4/10 = 0.40
Cloudy	2/10 = 0.20
Rainy	4/10 = 0.40

CPT for wind given weather P(wind | weather)

Weather	P(Weak)	P(Strong)
Sunny	3/4 = 0.75	1/4 = 0.25
Cloudy	1/2 = 0.50	1/2 = 0.50
Rainy	3/4 = 0.75	1/4 = 0.25

CPT for wind given weather P(humidity | weather)

Weather	P(High)	P(Normal)
Sunny	3/4 = 0.75	1/4 = 0.25
Cloudy	1/2 = 0.50	1/2 = 0.50
Rainy	1/4 = 0.25	3/4 = 0.75

Compare the naïve bayes which assumes indepe

P(W, Wi, H, T) = P(T) × P(W|T) × P(Wi|T) × P(H|T)

$$P(C_k \mid x_1, x_2, \ldots, x_n) = \frac{P(C_k) \times \prod_{i=1}^{n} P(x_i \mid C_k)}{P(x_1, x_2, \ldots, x_n)}$$

P(A|B) = P(A,B)/P(B)
P(PlayTennis | Weather, Wind, Humidity) = P(PlayTennis, Weather, Wind, Humidity)/ P(Weather, Wind, Humidity)

Note: The denominator is the same thing summed over both values of Play Tennis (Normalizer)

P(Weather) × P(Wind|Weather) × P(Humidity|Weather) × P(PlayTennis|Weather, Wind, Humidity) /
(P(Weather) × P(Wind|Weather) × P(Humidity|Weather) × P(Yes|Weather, Wind, Humidity) + P(Weather) × P(Wind|Weather) × P(Humidity|Weather) × P(No|Weather, Wind, Humidity))

Weather = Rainy, Wind = Weak, Humidity = High,
play= ?

Classify a new observation

Query: **Weather=Sunny, Wind=Weak, Humidity=Normal ?**

P(T | Sunny, Weak, Normal) = P(T, Sunny, Weak, Normal) / P(Sunny, Weak, Normal)

P(T, Sunny, Weak, Normal) = P(Sunny) × P(Weak|Sunny) × P(Normal|Sunny) × P(T|Sunny, Weak, Normal)

The denominator is obtained by summing the numerator over all values of T:

P(Sunny, Weak, Normal) = P(Sunny) × P(Weak|Sunny) × P(Normal|Sunny) × P(Yes|Sunny, Weak, Normal) +

P(Sunny) × P(Weak|Sunny) × P(Normal|Sunny) × P(No|Sunny, Weak, Normal)

Factor out the common terms:

P(Sunny, Weak, Normal) = P(Sunny) × P(Weak|Sunny) × P(Normal|Sunny) × [P(Yes|Sunny,Weak,Normal) + P(No|Sunny,Weak,Normal)]

Since P(Yes|Sunny,Weak,Normal) + P(No|Sunny,Weak,Normal) = 1:

P(Sunny, Weak, Normal) = P(Sunny) × P(Weak|Sunny) × P(Normal|Sunny)

P(T | Sunny, Weak, Normal) = P(Sunny) × P(Weak|Sunny) × P(Normal|Sunny) × P(T|Sunny, Weak, Normal)/P(Sunny) × P(Weak|Sunny) × P(Normal|Sunny)

The: P(Sunny) × P(Weak|Sunny) × P(Normal|Sunny) terms cancel and you are left with:

P(T | Sunny, Weak, Normal) = P(T | Sunny, Weak, Normal)

Weather	Wind	Humidity	P(Yes)	P(No)
Sunny	Weak	Normal	1/1 = 1.00	0/1 = 0.00